

# TISMO: syngeneic mouse tumor database to model tumor immunity and immunotherapy response

Zexian Zeng<sup>1,2,\*</sup>, Cheryl J. Wong<sup>1,3,†</sup>, Lin Yang<sup>1,†</sup>, Nofal Ouardaoui<sup>1</sup>, Dian Li<sup>1</sup>, Wubing Zhang<sup>1,4</sup>, Shengqing Gu<sup>5,6,1,2</sup>, Yi Zhang<sup>1,2</sup>, Yang Liu<sup>1</sup>, Xiaoqing Wang<sup>5,6,1,2</sup>, Jingxin Fu<sup>5,7</sup>, Liye Zhou<sup>5</sup>, Boning Zhang<sup>5,6,1,2</sup>, Sarah Kim<sup>7</sup>, Kathleen B. Yates<sup>7,8</sup>, Myles Brown<sup>5,6</sup>, Gordon J. Freeman<sup>5</sup>, Ravindra Uppaluri<sup>5,9</sup>, Robert Manguso<sup>7,8</sup> and X. Shirley Liu<sup>1,2,6,\*</sup>

<sup>1</sup>Department of Data Science, Dana Farber Cancer Institute, Boston, MA 02215, USA, <sup>2</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA, <sup>3</sup>Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115 USA, <sup>4</sup>School of Life Science and Technology, Tongji University, Shanghai, 200060, China, <sup>5</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA, <sup>6</sup>Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA 02215, USA, <sup>7</sup>Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02129, USA, <sup>8</sup>Center for Cancer Research, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02114, USA and <sup>9</sup>Department of Surgery, Brigham and Women's Hospital, Boston, MA 02215, USA

Received July 28, 2021; Revised August 28, 2021; Editorial Decision September 02, 2021; Accepted September 16, 2021

## ABSTRACT

Syngeneic mouse models are tumors derived from murine cancer cells engrafted on genetically identical mouse strains. They are widely used tools for studying tumor immunity and immunotherapy response in the context of a fully functional murine immune system. Large volumes of syngeneic mouse tumor expression profiles under different immunotherapy treatments have been generated, although a lack of systematic collection and analysis makes data reuse challenging. We present Tumor Immune Syngeneic MOuse (TISMO), a database with an extensive collection of syngeneic mouse model profiles with interactive visualization features. TISMO contains 605 *in vitro* RNA-seq samples from 49 syngeneic cancer cell lines across 23 cancer types, of which 195 underwent cytokine treatment. TISMO also includes 1518 *in vivo* RNA-seq samples from 68 syngeneic mouse tumor models across 19 cancer types, of which 832 were from immune checkpoint blockade (ICB) studies. We manually annotated the sample metadata, such as cell line, mouse strain, transplantation site, treatment, and response status, and uniformly processed and quality-controlled the RNA-seq data. Besides data download, TISMO provides in-

teractive web interfaces to investigate whether specific gene expression, pathway enrichment, or immune infiltration level is associated with differential immunotherapy response. TISMO is available at <http://tismo.cistrome.org>.

## INTRODUCTION

Immunotherapies targeting co-inhibitory pathways have shown remarkable clinical success but only demonstrate efficacy in a subset of cancer patients (1). The underlying mechanisms of heterogeneous response to immune checkpoint blockade (ICB) therapy remain unclear. Clinical samples from patient tumors accurately represent the tumor microenvironment (TME), but are difficult to access and conduct controlled experiments. Pre-clinical models that faithfully recapitulate the complexity of cancer cells and their interactions with the immune system are essential for investigating potential resistance mechanisms to ICB (2). Commonly used *in vitro* systems for cancer research, such as conventional 2D cell culture or 3D organoids, are inadequate to model the complexity of the TME. Instead, syngeneic tumors transplanted into immunocompetent mice are readily available and provide reproducible results for cancer immunology research. Syngeneic mouse models have been widely used in cancer immunology studies, and a large volume of tumor expression profiles under various immunotherapy treatments have been generated (3,4).

\*To whom correspondence should be addressed. Tel: +1 617 632 2472; Fax: +1 617 632 2444; Email: xslu@ds.dfci.harvard.edu  
Correspondence may also be addressed to Zexian Zeng. Tel: +1 617 632 6052; Fax: +1 617 632 6062; Email: zzeng@ds.dfci.harvard.edu  
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint first authors.

However, lack of systematic collection and variation in analysis of individually published syngeneic tumor profiles makes data reuse challenging.

Multiple existing data resources are available for mouse models of cancer, such as MPD (5,6), MMHCdb (7,8), MGD (9), GXD (10), PDX Finder (11) and NCI OMF (<https://oncologymodels.org/annotatedDataSets>). Among these databases, only NCI OMF contains syngeneic tumor model studies, although it solely provides meta-information of these studies without expression profiles. Similar to NCI OMF in providing study-level meta-information, GXD (10) focuses on expression profiles of wild-type and genetic mutant mice, but its scope is limited to embryonic stages and postnatal period. MPD (5,6) focuses on phenotypes of different mouse strains under specific experimental treatments, with strain-specific genotyping and microarray gene expression data for selected samples, although does not include syngeneic tumor models. MMHCdb (7,8) focuses on genetically engineered mouse models, inbred strains, and patient-derived xenograft models of human cancer and provides information about specific mutations/allelic variants in mouse tumors. MGD (9) is a major component of Mouse Genome Informatics (MGI) and provides descriptive annotations about mouse genes and other genome features such as nucleotide, protein sequences, and SNPs. Both MMHCdb (7,8) and MGD (9) explore the association between human diseases and mouse models by linking genetic background to phenotype, but neither include syngeneic tumor models. PDX Finder (11) is a searchable catalogue containing information for 1985 PDX models of diverse cancers, but as the name implies, this resource is limited to PDX models. To the best of our knowledge, there is no published database with a comprehensive collection of syngeneic mouse tumors that provides expression profiles and phenotypic data.

Herein, we present Tumor Immune Syngeneic MOuse (TISMO), a large-scale publicly accessible resource of syngeneic mouse models. TISMO (<http://tismo.cistrome.org>) is a comprehensive database with over two thousand uniformly processed and quality-controlled RNA-seq samples of syngeneic mouse cancer cell lines and tumor models. These datasets were uniformly processed from raw sequencing reads using a standardized workflow. In addition, immune cell infiltration and pathway enrichment levels have been inferred and phenotypic metadata have been manually annotated. TISMO provides interactive web interfaces for users to compare and visualize gene expression, pathway enrichment, and immune infiltration level across syngeneic mouse models, treatments, and response groups. The continued maintenance of TISMO will be of great utility to the cancer immunology and immuno-oncology research community.

## MATERIALS AND METHODS

### Data collection and meta information curation

We developed a parser to query datasets deposited in the Gene Expression Omnibus (GEO) (12) between 2016 and 2021. Using this parser, we performed keyword searches to identify studies matching a list of manually curated syngeneic mouse models or syngeneic cancer cell lines (Sup-

plementary Table S1). For matched studies, meta-files containing study design and sample information were downloaded through the parser. We manually curated and confirmed each sample for database inclusion. We also annotated syngeneic mouse model phenotypes through literature searches. In total, we collected 1868 syngeneic tumor or cell line RNA-seq samples from 137 published studies. We also included 255 in-house RNA-seq samples generated by ourselves which have not been published before.

### Transcriptome data processing

To ensure consistency, we downloaded raw sequencing reads from each study and processed the data through a standardized pipeline called RNA-seq IMMune Analysis Pipeline (RIMA, <https://liulab-dfci.github.io/RIMA>). RIMA is an automated Snakemake pipeline developed by our group to streamline the processing of RNA-seq data, including but not limited to read alignment, quality control, expression qualification, batch effect removal, and immune cell infiltration inference. FASTQ files containing the raw reads were downloaded or transferred. Read alignments were performed with STAR (13) (v.2.4.2a) against the mm10 reference genome assembly (mm10, Genome Reference Consortium Mouse Build 38) from the NCI Genomic Data Commons (GDC). RNA-seq quality control (QC) was performed on the aligned BAM files using RSeQC (14) (v2.4). With the reads appropriately aligned, expression levels were quantified by SALMON (15) (v.0.14.0) on the BAM files. Ensemble IDs were converted to mouse gene symbols (GRCm38.p6).

We also characterized pathway enrichment for each sample to enable comparison between conditions. From Molecular Signatures Database (MSigDB) (16), we first collected 17456 gene-sets, including 7479 GO biological processes, 996 GO cellular components, 1704 GO molecular functions (17), 186 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (18), 4872 immunologic signatures (16), 615 wiki pathways, and 1604 Reactome pathways (19). Then for each RNA-seq profile, we evaluated the level of each pathway by single sample gene set enrichment analysis (ssGSEA) (20). For user-defined gene set with weights, we calculate the weighted sum expression of the gene set following the equation of  $p = \sum_{i=1}^n k_i \log(e_i)$ , where  $p$  denotes the pathway level,  $k_i$  denotes the weights, and  $e_i$  denotes the transcripts per kilobase million (TPM) of the  $i$ th gene in the pathway.

Samples collected from different studies were processed with different protocols and platforms, and subject to technical bias. Even though TISMO conducts all sample comparisons within each cohort, we normalized the data for consistency across cohorts. Specifically, for each cohort we pre-computed false discovery rates (FDR) using DESeq2 (21) for each gene and comparison condition. To aid visualization, we normalized the transcriptome TPM data by quantile normalization to calibrate the scaling and distribution differences across samples, separately in each syngeneic model. We then performed batch effect correction between studies using ComBat within each syngeneic model (22). Notably, the combination of quantile normalization

and ComBat was reported to achieve the best performance for batch effect removal in a recent benchmark study (23). In an evaluation, we observed that after quantile normalization and batch effect removal, the samples' housekeeping-gene-signatures, averaged by 600 housekeeping genes (24), are less scattered between samples from different studies (Supplementary Figure S1A–D). The coefficient of variations reduced from 0.12 to 0.06 and from 0.14 to 0.09 for the *in vitro* samples and *in vivo* samples, respectively.

To make reliable and robust immune cell infiltration estimations, we utilized Immunedeconv (25), an R package that integrates state-of-the-art algorithms for immune deconvolution, including TIMER (26), xCell (27), CIBERSORT (28), EPIC (29), quanTIseq (30). We also incorporated a murine-specific immune infiltration deconvolution tool, mMCPcounter (31). Although each algorithm has unique properties and strengths (25), immune infiltration estimations supported by multiple algorithms provide more confident results. The source code to search mouse-related studies, download sequencing data, perform batch effect correction and infer immune infiltration were deposited at the Github repository (<https://github.com/zexian/TISMO.data>).

### Website development

To enable users to systematically explore the curated datasets, we developed a user-friendly web interface to host the TISMO database. The TISMO website is freely available at <http://tismo.cistrome.org> without any registration or login restriction. It is implemented with the R-Shiny framework (R version 3.6.3) on an Apache2 HTTP server, and is compatible with smartphones and tablets. The website consists of seven functional components: 'Home', 'Data Browser', 'Gene', 'Pathway', 'Infiltrates', 'Data Download', and 'Documentation'. 'Home' includes a tutorial video with step-by-step instructions on using the database and website. Users could browse or search the curated metadata using the 'Data Browser' module to locate relevant syngeneic models. 'Gene', 'Pathway' and 'Infiltrates' modules enable users to select and compare gene expression, pathway enrichment, and immune infiltration level between treatments, response groups, and models. Users could explore and derive gene expression programs or immune infiltrates consistently associated with ICB treatment and response in these modules. In the pathway module, users could also upload and evaluate self-defined gene sets. In the 'Data Download' module, users could download phenotypic metadata, quantified gene expression, and immune cell infiltration for all samples in the database. The documentation page summarizes the data processing steps and the number of samples in different models and treatment conditions.

## RESULTS

### Data summary

The current TISMO database includes 605 *in vitro* RNA-seq samples from 49 syngeneic cancer cell lines across 23 cancer types. TISMO also contains 1518 *in vivo* RNA-seq samples from 68 syngeneic mouse tumor models across 19 cancer types (Figure 1A and B). Many samples within

the TISMO database have undergone different treatments, including anti-PD1, anti-PDL1, anti-PDL2, anti-CTLA4, interferon  $\gamma$  (IFN $\gamma$ ), IFN $\beta$ , tumor necrosis factor alpha (TNF $\alpha$ ), or in combination with other treatments (Supplementary Tables S2 and S3). We manually annotated phenotypic data for each sample by referencing the original article, including cancer type, cancer cell line, cell treatment, cell genotype, mouse genotype, mouse strain, implantation type, implantation site, mouse ICB treatment, and response status (Supplementary Tables S2–S3). In addition, we have collected available survival information from published studies (Supplementary Figure S2). Gene expression levels were quantified for all samples, and immune cell infiltration levels for each *in vivo* sample were inferred based on expression profiles (Supplementary Table S4). All curated metadata, expression data, and immune infiltration estimation data can be downloaded from the 'Data Download' module in the TISMO database.

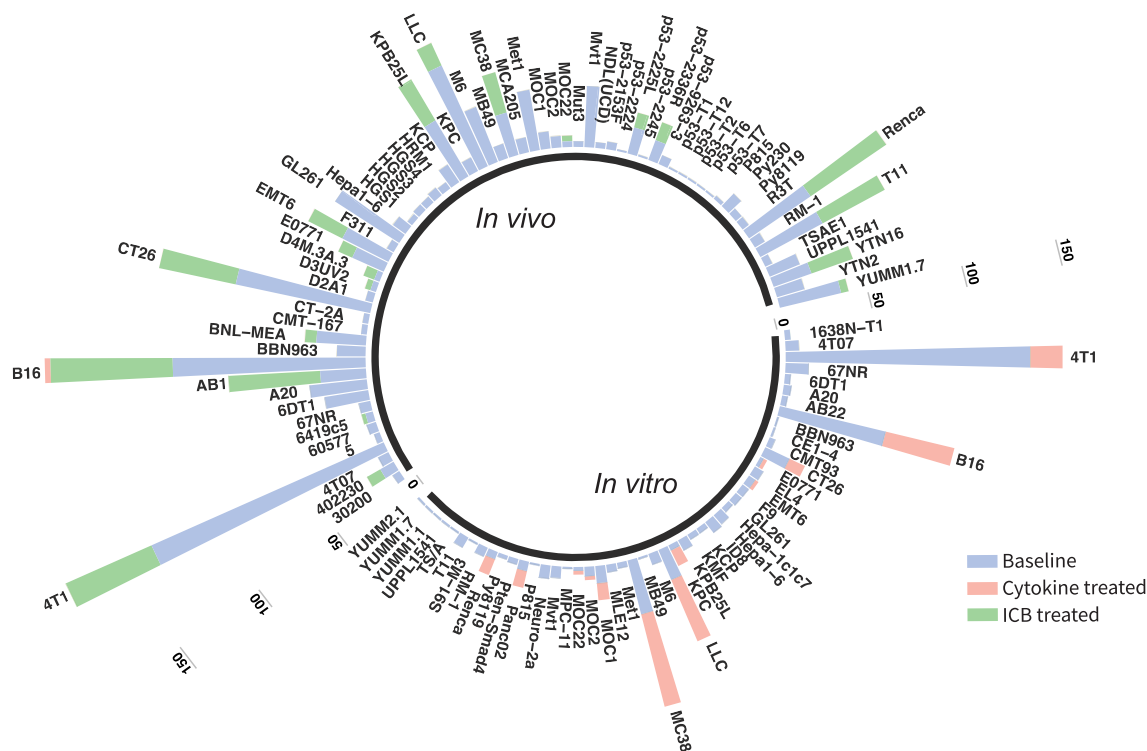
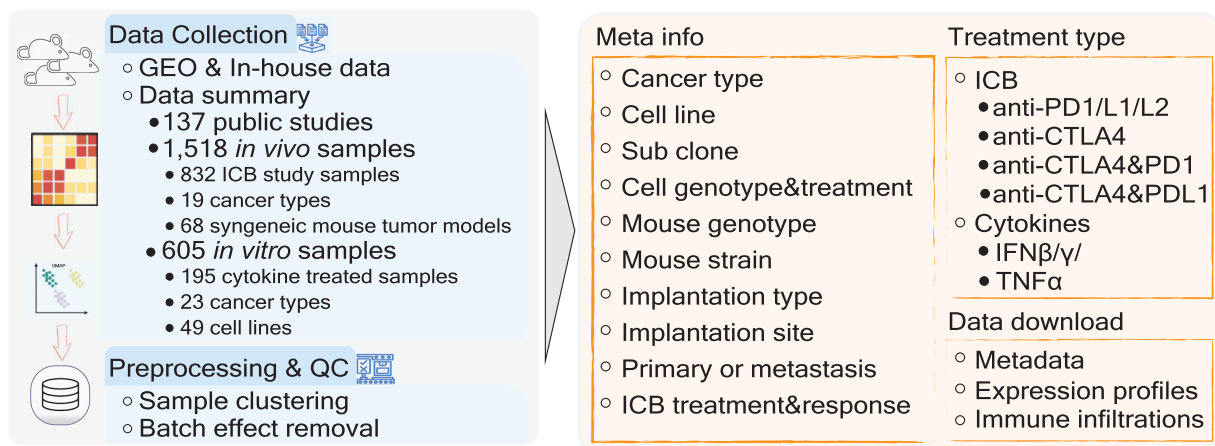
### Exploration of gene expression and pathway enrichment changes induced by cytokine treatment

Cytokines, including IFN and TNF, play essential roles in adaptive immunity in the TME (32–35). Due to their critical roles in anti-tumor immunity, it is of great interest to know how gene expression or pathway enrichment is differentially regulated by IFN or TNF stimulation. In TISMO, the *in vitro* data allow users to explore the effects of cytokine treatments on syngeneic cancer cell lines. After a user selects a gene or gene set, cytokine, and cell line, the TISMO web-server displays box plots of gene expression or pathway enrichment before and after cytokine treatment. For example, Figure 2A demonstrates how a user could investigate the 'MHC\_PROTEIN\_COMPLEX\_ASSEMBLY' pathway after cytokine stimulation in syngeneic cancer cell lines. After selecting the pathways, cytokines, and cell lines of interest, and submitting the query, a summary boxplot of the pathway level before and after cytokine treatment is generated together with statistical comparisons. As expected, there is a significant increase in major histocompatibility complex (MHC) protein complex after IFN $\gamma$  and IFN $\beta$  stimulation (32,36). In addition to the curated pathways, users could compare gene expression in the 'Gene' module or upload their own gene set of interest in the 'Upload pathway' module. The differentially expressed genes and pathway enrichment between comparison groups are statistically evaluated by the Wald test using DESeq2 (21) and the Student's t-test, respectively. In TISMO, we have curated 17456 pathways from MSigDB (16) and have characterized pathway enrichment for each sample using ssGSEA (20), which users can explore interactively through the web interface. Users could also upload self-defined gene sets with the flexibility to adjust these gene weights. If a user uploads a customized gene set, TISMO will calculate its level in each *in vitro* and *in vivo* sample (Materials and Methods), allowing users to evaluate them across models, treatments, and response groups.

### Comparison of gene expression, pathway enrichment, and immune infiltration between ICB treatment and response groups

ICB treatments, including anti-CTLA4 and anti-PD1/L1 provide clinical benefits in only a subset of patients (37).

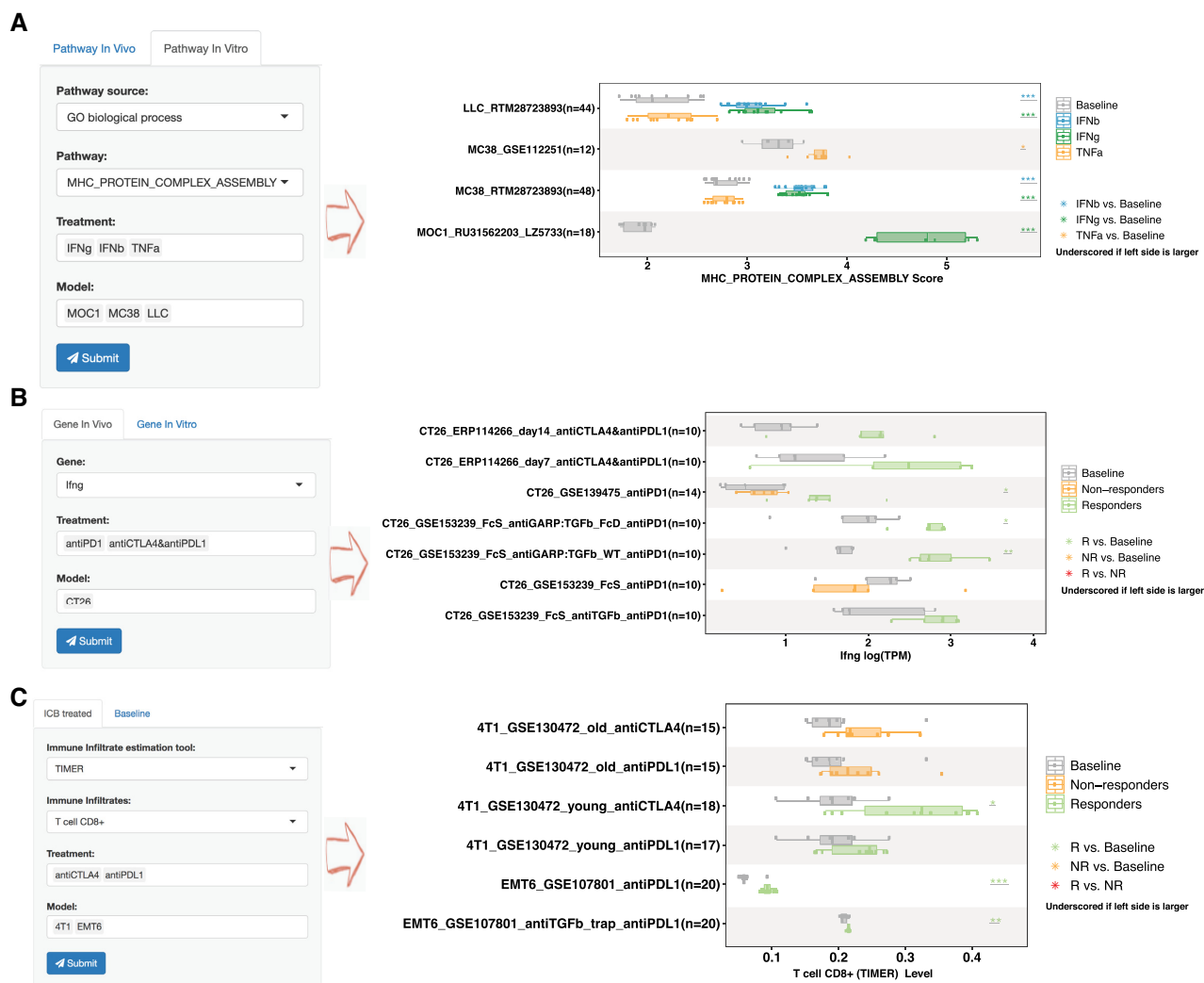




**Figure 1.** Overview of the TISMO database and summary of data available for download. **(A)** The datasets hosted at the TISMO database were curated from public domains and in-house data. In total, TISMO contains 605 *in vitro* samples, of which 195 were cytokine treated. TISMO also contains 1518 *in vivo* samples, of which 832 samples were from immune checkpoint blockade (ICB) studies. All samples were uniformly processed from raw sequencing reads within a standardized workflow, including alignment, quality control, expression qualification, batch effect removal, and immune cell infiltration inference. In addition, sample metadata, including cell line, mouse strain, transplantation site, ICB treatment, survival and response status were manually annotated from publications. The metadata, expression profiles, and immune cell expression profiles for all samples can be downloaded from the TISMO database. **(B)** TISMO contains *in vitro* RNA-seq samples from 49 syngeneic cancer cell lines across 23 cancer types. TISMO also contains *in vivo* samples from 68 syngeneic mouse tumor models across 19 cancer types. Samples treated with cytokines (IFN $\gamma$ , IFN $\beta$ , TNF $\alpha$ ) are labeled in pink; samples treated with ICB (anti-PD1, anti-PDL1, anti-PDL2, antiCTLA4 or their combinations) are labelled in green.

Mechanisms underlying heterogeneous ICB response remain an open question and the association between immune infiltration and ICB response is of great interest to many researchers (26). In addition to gene expression profiles and pathway enrichment, the TISMO database integrates inferred immune infiltrations from six state-of-the-art immune cell deconvolution algorithms, including

TIMER (26), xCell (27), CIBERSORT (28), EPIC (29), quanTIseq (30) and mMCPcounter (31). We have annotated treatment and response information for ICB study samples ( $N = 832$ ), allowing users to explore whether specific gene expression programs or immune infiltrations are robustly associated with ICB response or resistance. Using TISMO, users could select a specific gene, pathway, or



**Figure 2.** Examples of TISMO features exploring factors associated with cytokine treatment and immunotherapy response. (A) Users could select pathways, cytokines, and syngeneic cell lines to evaluate their expressions before and after cytokine treatment. (B, C) Users could select a gene (B) and immune cell infiltration (C) to evaluate whether its level changes upon ICB treatment, and compare it between ICB responders and non-responders. Immune cell infiltration levels were inferred by six different state-of-the-art deconvolution algorithms. The generated figures could be downloaded in jpg, pdf, and table formats.

immune cell infiltration, evaluate whether its level changes upon ICB treatment, and compare it between ICB responders and non-responders (Figure 2B, C). For example, Figure 2B shows how the IFN $\gamma$  gene expression in CT26 model is stimulated by different ICB treatments. After a user selects genes, ICB treatments, and syngeneic mouse models of interest, TISMO website will generate a box plot summarizing the gene expression levels before and after ICB treatments in different comparison conditions. The expression levels are grouped by syngeneic model, treatment group, and response status in the figure. After ICB treatment, the IFN $\gamma$  expression levels in CT26 model are significantly upregulated in the ICB responders, but not in the non-responders (Figure 2B) (32). In addition to gene expression levels, users could also compare pathway enrichment levels in the ‘Pathway’ module or immune cell infiltration in the ‘Infiltrate’ module (Figure 2C). The differential gene expression between groups is statistically evaluated by the Wald test in DESeq2 (21), and the differential pathway enrich-

ment (characterized by ssGSEA) (20)) and immune infiltration are evaluated by the Mann–Whitney  $U$  test. Compared to clinical samples, syngeneic tumor models allow scientists to evaluate immunotherapy response in a more controlled and reproducible manner. TISMO’s website enables users to efficiently evaluate genes, pathways, and immune cell infiltration in the context of ICB treatment, to generate or validate hypotheses on immunotherapy response.

### Exploration of meta information, figure generation and data download

TISMO’s website hosts a data browser module to help users locate relevant mouse models from our collection. Users could query sample Study ID (the majority representing the GSE ID, a study identification number for the GEO database), sample metadata, treatment condition, response status and the number of replicates in each design. The data browser module aids researchers in selecting the most rel-

evant syngeneic mouse models to supply evidence for hypothesis generation or validation. On the TISMO website, data cohorts and treatment types could be selected for statistical comparison and visualization. The interactive filtering and visualization features enable users to systematically compare different models, cell lines, treatments, and response groups. The generated figures could be downloaded in jpg, pdf, and table formats. The download page provides access to the expression matrices of all 2123 RNA-seq profiles, immune cell infiltration estimations of the *in vivo* samples, and the manually annotated metadata for all samples.

## DISCUSSION

Syngeneic mouse models are essential in immunotherapy research as they enable the study of cancer cells in the context of immunocompetent hosts. Large amounts of syngeneic mouse model profiles have been generated, but these data are scattered, making data reuse challenging. There is still no comprehensive, intuitive, and convenient database with user-friendly, interactive web features for researchers to explore syngeneic mouse model data. TISMO (<http://tismo.cistrome.org>) is the first comprehensive database for users to investigate and visualize gene expression, pathway enrichment, and immune cell infiltration levels in syngeneic mouse models across different ICB treatment and response groups. Expression profiles hosted on TISMO were uniformly processed from the raw sequencing reads. The immune infiltration levels were inferred by six state-of-the-art deconvolution algorithms, and the metadata were manually annotated from publications. TISMO provides web interfaces to help users explore the syngeneic mouse model data interactively. In summary, TISMO is a comprehensive database of syngeneic mouse models that will help users select relevant syngeneic mouse tumor models, provide data to generate and test hypotheses, and reveal novel mechanisms of ICB response and resistance.

## DATA AVAILABILITY

TISMO is available at <http://tismo.cistrome.org> to all users without restrictions. Annotated mouse syngeneic metadata are available on the website. All expression data and immune infiltration estimation can be downloaded from the data download page.

## SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

## ACKNOWLEDGEMENTS

We acknowledge authors from published studies for sharing their data.

## FUNDING

Breast Cancer Research Foundation [BCRF-20-100]; National Institutes of Health [R01CA234018, U24CA224316, T15LM007092]; Sara Elizabeth O'Brien Trust; Dana-Farber Cancer Institute. Funding for open access charge: National Institutes of Health [R01CA234018].

**Conflict of interest statement.** M.B. is a consultant to and receives sponsored research support from Novartis. M.B. is a consultant to and serves on the scientific advisory boards of Kronos Bio, H3 Biomedicine, and GV20 Oncotherapy. G.J.F. has patents/pending royalties on the PD-1/PD-L1 pathway from Roche, Merck MSD, Bristol-Myers-Squibb, Merck KGA, Boehringer-Ingelheim, AstraZeneca, Dako, Leica, Mayo Clinic and Novartis. G.J.F. has served on advisory boards for Roche, Bristol-Myers-Squibb, Xios, Origimed, Triursus, iTeos, NextPoint, IgM, Jubilant, Trilium, GV20 and Geode. G.J.F. has equity in Nextpoint, Triursus, Xios, iTeos, IgM, GV20 and Geode. R. Uppaluri is a consultant for and has received research support from Merck. R.T.M. consults for Bristol-Myers Squibb and Tango Therapeutics. X.S.L. is a cofounder, board member, SAB member, and consultant of GV20 Oncotherapy and its subsidiaries. X.S.L. is stockholder of BMY, TMO, WBA, ABT, ABBV and JNJ, and received research funding from Takeda, Sanofi, Bristol Myers Squibb, and Novartis. No potential conflicts of interest were disclosed by the other authors.

## REFERENCES

- Waldman, A.D., Fritz, J.M. and Lenardo, M.J. (2020) A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat. Rev. Immunol.*, **20**, 651–668.
- Day, C.-P., Merlino, G. and Van Dyke, T. (2015) Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell*, **163**, 39–53.
- Olson, B., Li, Y., Lin, Y., Liu, E.T. and Patnaik, A. (2018) Mouse models for cancer immunotherapy research. *Cancer Discov.*, **8**, 1358–1365.
- Chulpanova, D.S., Kitaeva, K.V., Rutland, C.S., Rizvanov, A.A. and Solovyeva, V.V. (2020) Mouse tumor models for advanced cancer immunotherapy. *Int. J. Mol. Sci.*, **21**, 4118.
- Grubb, S.C., Bult, C.J. and Bogue, M.A. (2014) Mouse phenome database. *Nucleic Acids Res.*, **42**, D825–D834.
- Bogue, M.A., Philip, V.M., Walton, D.O., Grubb, S.C., Dunn, M.H., Kolishovski, G., Emerson, J., Mukherjee, G., Stearns, T., He, H. *et al.* (2020) Mouse Phenome Database: a data repository and analysis suite for curated primary mouse phenotype data. *Nucleic Acids Res.*, **48**, D716–D723.
- Begley, D.A., Krupke, D.M., Vincent, M.J., Sundberg, J.P., Bult, C.J. and Eppig, J.T. (2007) Mouse Tumor Biology Database (MTB): status update and future directions. *Nucleic Acids Res.*, **35**, D638–D642.
- Bult, C.J., Krupke, D.M., Begley, D.A., Richardson, J.E., Neuhauser, S.B., Sundberg, J.P. and Eppig, J.T. (2015) Mouse Tumor Biology (MTB): a database of mouse models for human cancer. *Nucleic Acids Res.*, **43**, D818–D824.
- Bult, C.J., Kadin, J.A., Richardson, J.E., Blake, J.A., Eppig, J.T. and Mouse Genome Database, G. (2010) The mouse genome database: enhancements and updates. *Nucleic Acids Res.*, **38**, D586–D592.
- Baldarelli, R.M., Smith, C.M., Finger, J.H., Hayamizu, T.F., McCright, I.J., Xu, J., Shaw, D.R., Beal, J.S., Blodgett, O. and Campbell, J. (2021) The mouse Gene Expression Database (GXD): 2021 update. *Nucleic Acids Res.*, **49**, D924–D931.
- Conte, N., Mason, J.C., Halmagyi, C., Neuhauser, S., Mosaku, A., Yordanova, G., Chatzipili, A., Begley, D.A., Krupke, D.M. and Parkinson, H. (2019) PDX finder: a portal for patient-derived tumor xenograft model discovery. *Nucleic Acids Res.*, **47**, D1073–D1079.
- Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M. and Holko, M. (2012) NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.*, **41**, D991–D995.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
- Wang, L., Wang, S. and Li, W. (2012) RSeQC: quality control of RNA-seq experiments. *Bioinformatics*, **28**, 2184–2185.

15. Patro,R., Duggal,G., Love,M.I., Irizarry,R.A. and Kingsford,C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*, **14**, 417.
16. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R. and Lander,E.S. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15545–15550.
17. Gene Ontology Consortium (2021) The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res.*, **49**, D325–D334.
18. Kanehisa,M., Furumichi,M., Sato,Y., Ishiguro-Watanabe,M. and Tanabe,M. (2021) KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.*, **49**, D545–D551.
19. Jassal,B., Matthews,L., Viteri,G., Gong,C., Lorente,P., Fabregat,A., Sidiropoulos,K., Cook,J., Gillespie,M. and Haw,R. (2020) The reactome pathway knowledgebase. *Nucleic Acids Res.*, **48**, D498–D503.
20. Barbie,D.A., Tamayo,P., Boehm,J.S., Kim,S.Y., Moody,S.E., Dunn,I.F., Schinzel,A.C., Sandy,P., Meylan,E. and Scholl,C. (2009) Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*, **462**, 108–112.
21. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
22. Johnson,W.E., Li,C. and Rabinovic,A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**, 118–127.
23. Müller,C., Schillert,A., Röhmeier,C., Trégouët,D.-A., Proust,C., Binder,H., Pfeiffer,N., Beutel,M., Lackner,K.J. and Schnabel,R.B. (2016) Removing batch effects from longitudinal gene expression-quantile normalization plus ComBat as best approach for microarray transcriptome data. *PLoS One*, **11**, e0156594.
24. Hounkpe,B.W., Chenou,F., de Lima,F. and De Paula,E.V. (2021) HRT Atlas v1. 0 database: redefining human and mouse housekeeping genes and candidate reference transcripts by mining massive RNA-seq datasets. *Nucleic Acids Res.*, **49**, D947–D955.
25. Sturm,G., Finotello,F. and List,M. (2020) Immunedeconv: an R Package for Unified Access to Computational Methods for Estimating Immune Cell Fractions from Bulk RNA-Sequencing Data. *Methods Mol. Biol.*, **2120**, 223–232.
26. Li,T., Fu,J., Zeng,Z., Cohen,D., Li,J., Chen,Q., Li,B. and Liu,X.S. (2020) TIMER2. 0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.*, **48**, W509–W514.
27. Aran,D., Hu,Z. and Butte,A.J. (2017) xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.*, **18**, 220.
28. Newman,A.M., Liu,C.L., Green,M.R., Gentles,A.J., Feng,W., Xu,Y., Hoang,C.D., Diehn,M. and Alizadeh,A.A. (2015) Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods*, **12**, 453–457.
29. Racle,J., de Jonge,K., Baumgaertner,P., Speiser,D.E. and Gfeller,D. (2017) Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife*, **6**, e26476.
30. Finotello,F., Mayer,C., Plattner,C., Laschober,G., Rieder,D., Hackl,H., Krogstad,A., Loncova,Z., Posch,W., Wilflingseder,D. et al. (2019) Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.*, **11**, 34.
31. Petitprez,F., Levy,S., Sun,C.M., Meylan,M., Linhard,C., Becht,E., Elarouci,N., Tavel,D., Roumenina,L.T., Ayadi,M. et al. (2020) The murine Microenvironment Cell Population counter method to estimate abundance of tissue-infiltrating immune and stromal cell populations in murine samples using gene expression. *Genome medicine*, **12**, 86.
32. Benci,J.L., Johnson,L.R., Choa,R., Xu,Y., Qiu,J., Zhou,Z., Xu,B., Ye,D., Nathanson,K.L., June,C.H. et al. (2019) Opposing functions of interferon coordinate adaptive and innate immune responses to cancer immune checkpoint blockade. *Cell*, **178**, 933–948.
33. Vredevoogd,D.W., Kuilman,T., Ligtenberg,M.A., Boshuizen,J., Stecker,K.E., de Bruijn,B., Krijgsman,O., Huang,X., Kenski,J.C.N., Lacroix,R. et al. (2019) Augmenting immunotherapy impact by lowering tumor TNF cytotoxicity threshold. *Cell*, **178**, 585–599.
34. Kearney,C.J., Vervoort,S.J., Hogg,S.J., Ramsbottom,K.M., Freeman,A.J., Lalaoui,N., Pijpers,L., Michie,J., Brown,K.K., Knight,D.A. et al. (2018) Tumor immune evasion arises through loss of TNF sensitivity. *Sci. Immunol.*, **3**, eaar3451.
35. Ayers,M., Lunceford,J., Nebozhyn,M., Murphy,E., Loboda,A., Kaufman,D.R., Albright,A., Cheng,J.D., Kang,S.P., Shankaran,V. et al. (2017) IFN- $\gamma$ -related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.*, **127**, 2930–2940.
36. Früh,K. and Yang,Y. (1999) Antigen presentation by MHC class I and its regulation by interferon  $\gamma$ . *Curr. Opin. Immunol.*, **11**, 76–81.
37. Patel,S.J., Sanjana,N.E., Kishton,R.J., Eidizadeh,A., Vodnala,S.K., Cam,M., Gartner,J.J., Jia,L., Steinberg,S.M., Yamamoto,T.N. et al. (2017) Identification of essential genes for cancer immunotherapy. *Nature*, **548**, 537–542.